

Assessment of Membrane Bioreactor Fouling Behaviour using Principal Component Analysis

W. Naessens*, T. Maere*, K. Villez**, S. Marsili-Libelli***, I. Nopens*

* BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, B-9000, Ghent, Belgium (Email: wouter.naessens@ugent.be; thomas.maere@ugent.be; ingmar.nopens@ugent.be)

** Laboratory of Intelligent Process Systems, School of Chemical Engineering, Purdue University, West Lafayette, IN, USA (Email: kvillez@purdue.edu)

*** Department of Systems and Computers, University of Florence, Via S. Marta 3, 50139 Florence, Italy (Email: marsili@dsi.unifi.it)

Abstract

While membrane bioreactors have become state-of-the-art technology, fouling is still responsible for high operational costs. Mechanistic modelling is hampered by the complexity of the fouling mechanisms. Therefore, principal component analysis and clustering are used in this contribution to extract information on the membrane state from both lab and pilot scale data sets. It was concluded from all analyses that the first principal component can be used for fouling monitoring, while the second principal component differentiates between reversible and irreversible fouling, if inter-membrane variance was included in the data set. On pilot scale, a low measuring frequency hampered clear results, but also here PCA-based techniques have potential for fouling monitoring.

Keywords

Membrane fouling; monitoring; principal component analysis

INTRODUCTION

Membrane bioreactors (MBR), being the coupling of a bioreactor for nutrient removal and membrane units for sludge retention, are considered state-of-the-art technology for wastewater treatment (Drews, 2010). They have several advantages over conventional activated sludge (CAS) systems. MBRs can be built on smaller sites thanks to the elimination of large secondary settling tanks. Next to that, the effluent is of better quality, not least by the retention of persistent chemicals, viruses and pathogens. The major drawback of MBRs is fouling of the membranes by particulates, colloids and solutes, increasing the operational costs (membrane aeration, cleaning actions, membrane replacements). To mitigate fouling problems, a lot of mechanistic fouling models have been built. However, since the underlying fouling mechanisms interact in ways that are at the moment not fully understood, a lot of contradictory results are reported. For data-driven modelling, a lot of data are usually available and a complete understanding of the fouling mechanism is not of vital importance.

In this contribution, principal component analysis (PCA), a data mining technique, is used in combination with a clustering algorithm to monitor fouling severity, discriminating between different fouling types, and thereby indicating the current membrane state. It allows automating the monitoring of different parameters at the same time, recognising data patterns and revealing the hidden relations within the data. This technique has already proved itself useful for fault detection in conventional wastewater treatment plants (Aguado and Rosén, 2008; Villez et al., 2008; Baggiani and Marsili-Libelli, 2009), but the application to MBR data has not yet been reported.

MATERIALS AND METHODS

Principal component analysis

Principal component analysis (PCA) is a data mining technique, introduced by Pearson (1901) and

Hotelling (1933). It is a mathematical tool providing a data reduction and the efficient extraction of information from large data sets. The method consists of several steps, the first being a transformation of the measured variables $X_{.,i}$ into a set of new variables $T_{.,i}$ called principal components (PC), thereby eliminating correlation between the variables. This is done by making linear combinations of the original variables, in which the weight vectors $P_{.,i}$ are called loading vectors.

$$X * P = T \quad (1)$$

The determination of the loading vectors is key to the data reduction. The calculation implies the maximization of the variability in the corresponding principal component, which has to be orthogonal to all other principal components, while the loading vector itself is being restricted to unit norm.

$$\max_{P_{.,j}^T * P_{.,j} = 1, \text{cov}(T_{.,i}, T_{.,j}) = 0} \left(\text{var}(T_{.,j}) \right) \quad \forall i < j \quad (2)$$

As such, PCA can be seen as an axis transformation from a space formed by the original variables, to an orthonormal basis oriented according to the largest variance present in the data. It has been proven that the loading vectors following those restrictions can be found as the eigenvectors of the covariance matrix S of X , while the eigenvalues are equal to the variance in the principal components. Therefore, the next step in PCA is a data reduction, performed by selecting only the principal components corresponding to the highest eigenvalues, thus retaining a maximal amount of information (i.e. maximum variance) in a minimal amount of variables.

Functional principal component analysis

In functional principal component analysis (FPCA), the data are assumed to be discrete measurements of a continuous function. The data are therefore translated to a functional form. The traditional way to do this uses a set of basis functions, from which linear combinations are made to approximate the underlying function in a least-square sense. The weights allocated to the basis functions then form the functional data matrix for the PCA. Mostly, the interpretation of these data is not very straightforward. An alternative way, used in this contribution, is defining a number of parameters by which the function can be described accurately. If chosen well, a useful interpretation of the PCA results can be achieved. In this work, five parameters ($+\Delta P$, $-\Delta P$, a , b and S) were estimated to describe a typical TMP cycle, as given in Figure 1. The reasoning behind this choice and the estimation procedures are described later in the manuscript.

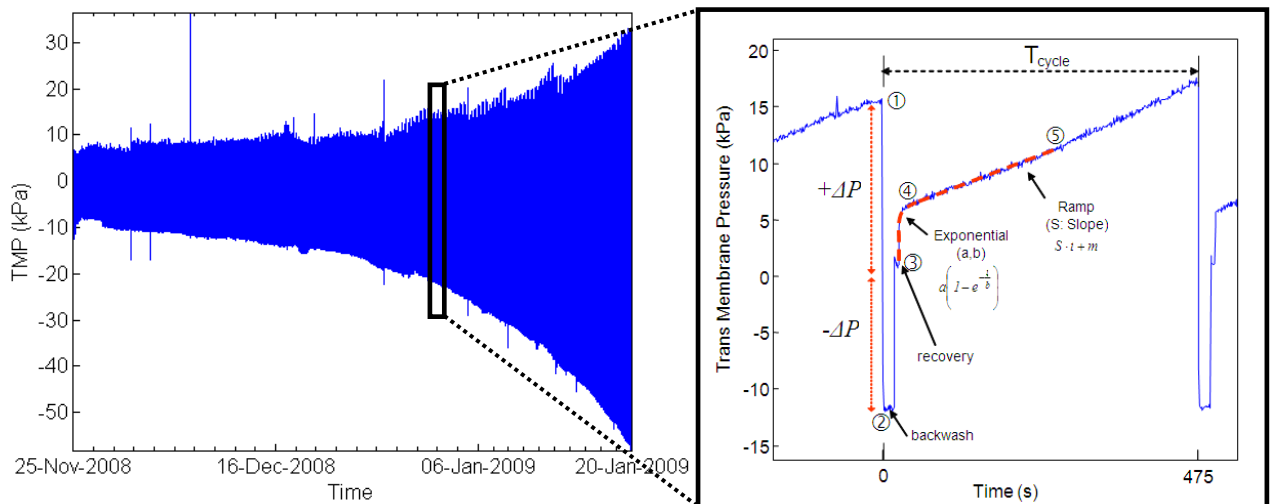


Figure 1. Left: example of a TMP profile in time (2 months of data). Right: detail of a typical filtration cycle and physical appearance and meaning of the estimated parameters.

Data sets

In this explorative study, transmembrane pressure (TMP) data from both a lab-scale and pilot scale MBR were used. The lab-scale installation (28.8L) is a sidestream configuration (X-Flow, Pentair) and is located at the BIOMATH research group (Ghent University, Belgium). The system is described in detail elsewhere (Jiang, 2008). The pilot scale installation is also a sidestream configuration and is located at the wastewater treatment plant of Ootmarsum (The Netherlands) and property of the water board Regge en Dinkel. Some operating parameters are given in Table 1, to provide an idea of how both systems were operated.

Table 1. Operating parameter values of the pilot scale MBR in Ootmarsum and the lab-scale MBR at BIOMATH.

Parameter (units)	Pilot scale	Lab-scale
Flux (L/m ² h)	45-60	31.8
Circulation flow (m ³ /h)	20	0.46
Airlift flow (Nm ³ /h)	10	0.46
Backpulse flow (L/m ² h)	350	106
Interval (filtr./backw./relax.)	7 min / 7s / -	7.5 min / 18s / 7s
MLSS (g/L)	10-14	8-10

Data processing

All data processing steps are visualized in Figure 2. Before conducting a PCA or FPCA, the data were unfolded per cycle in an automated way, using the first and second derivative of the TMP to locate the start of the backwash phase (Figure 1). As such, all measurements of the same cycle are treated as one multivariate sample, so later on an evaluation can be made for each cycle. For the PCA pathway, these unfolded data are used as input. Since every variable has the same unit (kPa), a data standardization process was not considered necessary and therefore not performed. For FPCA, five parameters were deduced from a typical TMP cycle, as shown in Figure 1. The pressure at the end of filtration, denoted by $+\Delta P$, gives an indication of fouling in general, while the negative pressure $-\Delta P$ resembles the backwash phase (mean value) and thus by definition only irreversible fouling. As for the filtration itself, the curve consists of an exponential part at the beginning of filtration and a subsequent linear part. Hence, the parameters “a” and “b” of an exponential model and the slope “S” of a linear model were estimated, using a least squared error curve fitting approach. Parameter “a” will be influenced by both reversible and irreversible fouling, while “b” is generally not used in literature, and S is related to reversible fouling only. All parameters were smoothened using the MATLAB cubic smoothing spline function. In contrast with common PCA, an autoscaling process was used to remove artificial differences in importance imposed by the strongly different units of the parameters.

After calculating the loading vectors and principal components, retaining only the PCs with the largest eigenvalues, thereby retaining a minimum of 95% (arbitrary) of the present variance, leads to a data reduction. Principal component values, called principal scores, were then clustered using the Gustafson-Kessel algorithm (Gustafson and Kessel, 1979; Babuska et al., 1998). Three cluster centers were chosen to obtain a separate cluster for clean, transitional and fouled behaviour. A reconstruction of the original variables was performed, thereby also treating the cluster centers as data points to obtain their position in the original coordinates. Based mainly on the clustering and the reconstruction plot, an assessment was made on the membrane state. All calculations were performed using MATLAB (Math Works, USA).

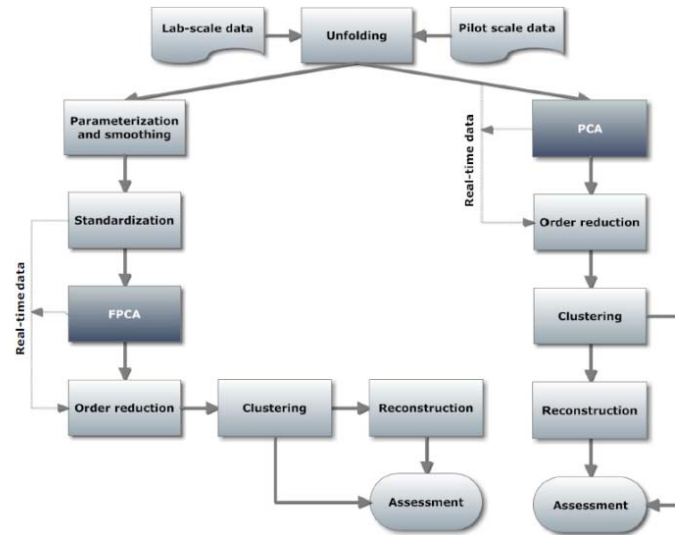


Figure 2. Flowchart of all data processing steps. In the future, real-time data can skip the model building step, and use a model structure built on historical data.

RESULTS AND DISCUSSION

Short term models for lab-scale reactor

As for the lab-scale installation, when fouling became problematic according to the expertise of the operator, the membrane module was replaced instead of using chemical cleanings. To gain insight in the structure of the data set, three different models were built, each for a time span of exactly the life time of one membrane module. As such, only intra-membrane variance was investigated and a comparison between the different models is possible. The used TMP data are given in Figure 3.

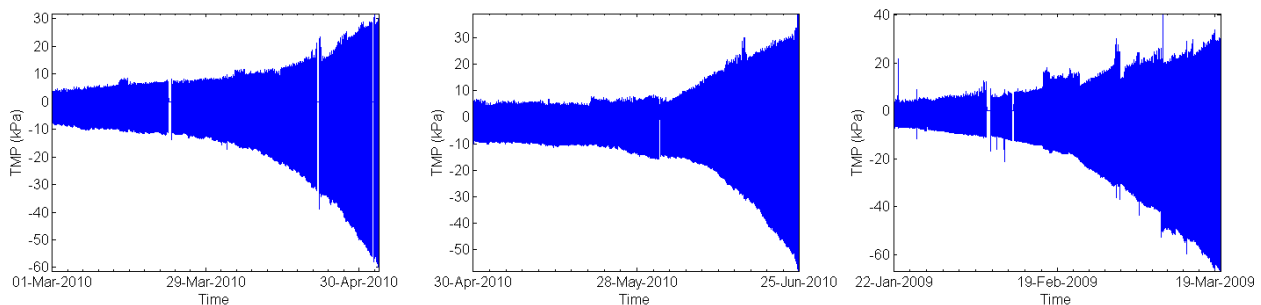


Figure 3. TMP time series for membrane 1 (M1), membrane 2 (M2) and membrane 3 (M3).

The following results are originating from the FPCA on these data sets. For the first membrane (M1) two PCs were needed to capture the arbitrary 95% of the originally present variance (not shown). A biplot of these first two components (Figure 4) reveals the structure of the model. For PC₁, all parameters are almost equally important, but there is a contrast between $-\Delta P$ and the other parameters. Since high positive values for $+\Delta P$, a , b and S combined with negative values for $-\Delta P$ typically result in a TMP cycle which is characteristic for a highly fouled membrane, PC₁ can be seen as a good indicator for fouling severity. For PC₂, b is of major importance, which could not be linked to a certain fouling condition, since the response of this parameter on different fouling conditions is not yet reported. For the second investigated membrane (M2), operated subsequently to the first (hence, exposed to the same sludge), the results were very similar: again two PCs were withheld based on retaining 95% variation and their compositions were very similar. However, for a third membrane (M3), operated more than one year earlier, the results differ considerably, especially in the composition of PC₂: now S is of major importance.

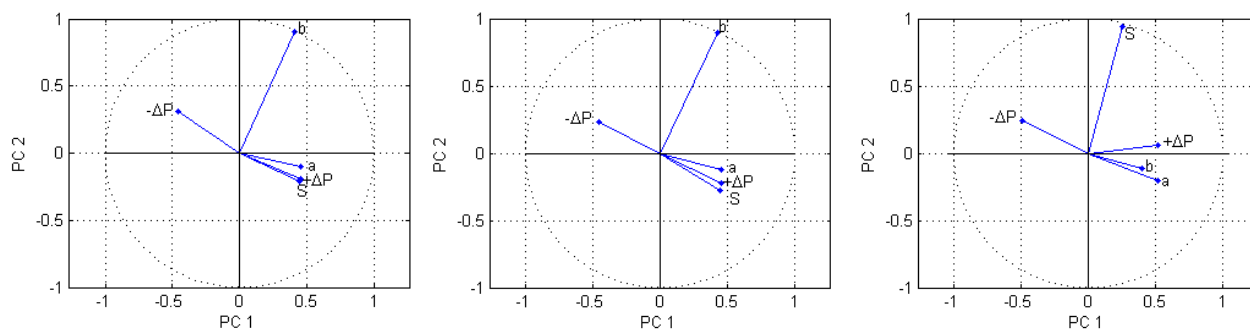


Figure 4. Biplot for membrane 1 (M1), membrane 2 (M2) and membrane 3 (M3). The loadings of the first two loading vectors, corresponding to the first two PCs, are drawn as coordinates. The construction of a PC out of the original variables is found by projecting all lines orthogonally onto the correct axis. Correlations between the original variables correspond to the angles between them (a small, opposite or perpendicular angle equals positive, negative and no correlation respectively.).

Since the MBR was operated in steady state, with a synthetic influent of constant composition, the change in model structure can be explained by a different equilibrium state of the system, being responsible for a different fouling mechanism. This change in mechanism alters the TMP profile subtly and thus the variations in the deduced parameters, eventually leading to a different model construction.

Long term model for lab-scale reactor

In an attempt to also capture inter-membrane differences within the principal components, rather than ending up with different model structures for each membrane, also a long term data set was analysed, with data of three consecutively used membranes. Note that the data of the second membrane of this data set are the same data as for the third analysed membrane in the short term analysis (M3). The complete data set is visualized in Figure 5.

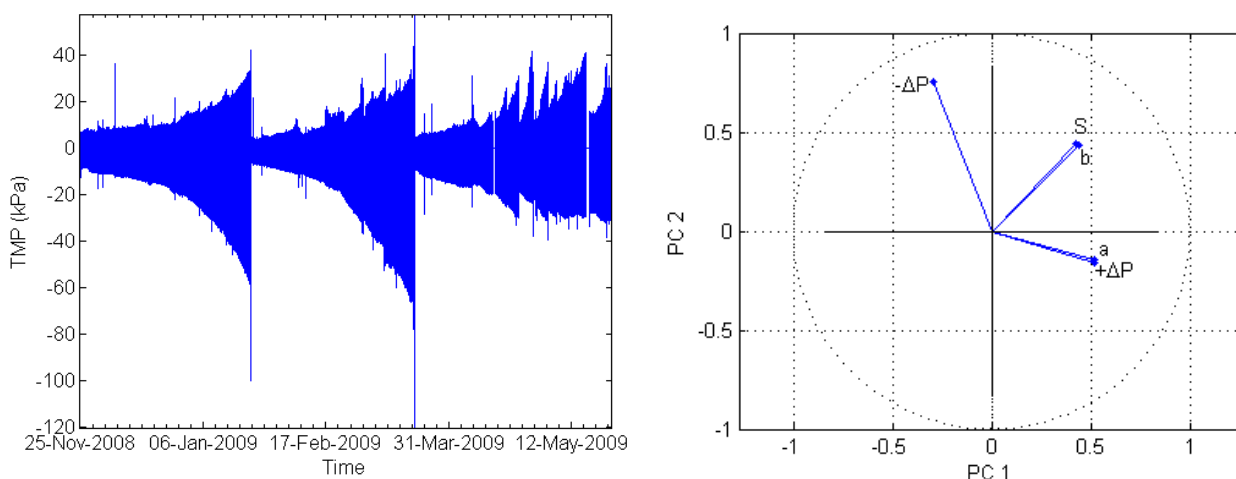


Figure 5. Left: TMP data set of three consecutively used membranes to form a longer term data set. Right: biplot for the first two PCs of the long term data set.

According to the biplot (Figure 5), the contrast in the first principal component is similar to the short term models. This implies that PC₁ is again a measure for fouling severity. However, the second principal component appears to be very different. A contrast can be seen between +ΔP and a on one hand and -ΔP, S and b on the other, but especially -ΔP, S and b are of importance. It was hypothesized that high scores for PC₂ thus correspond mainly to reversible fouling, characterized by less negative values for -ΔP and high values for S, while lower scores for PC₂ can be related to irreversible fouling using the opposite reasoning.

Figure 6 shows a cluster plot of the principal scores for the first two principal components, and a clear distinction can be made between the third membrane and the first two. According to the theory above, membrane 3 suffered mainly from reversible fouling, while irreversible fouling was the dominant fouling type of membrane 1 and 2. This is confirmed by the observation that the scores of PC_1 for membrane 3 are also higher, since S has a higher absolute contribution in the loadings of PC_1 compared to $-\Delta P$, indicating that reversible fouling leads to higher scores for PC_1 . From the reconstruction of the parameters (Figure 6), it can be seen that for the third membrane the fouled cluster for $-\Delta P$ is never reached, while it is the only membrane reaching the fouled cluster for S . Also these findings confirm the hypothesis that reversible fouling was only dominant for the third membrane.

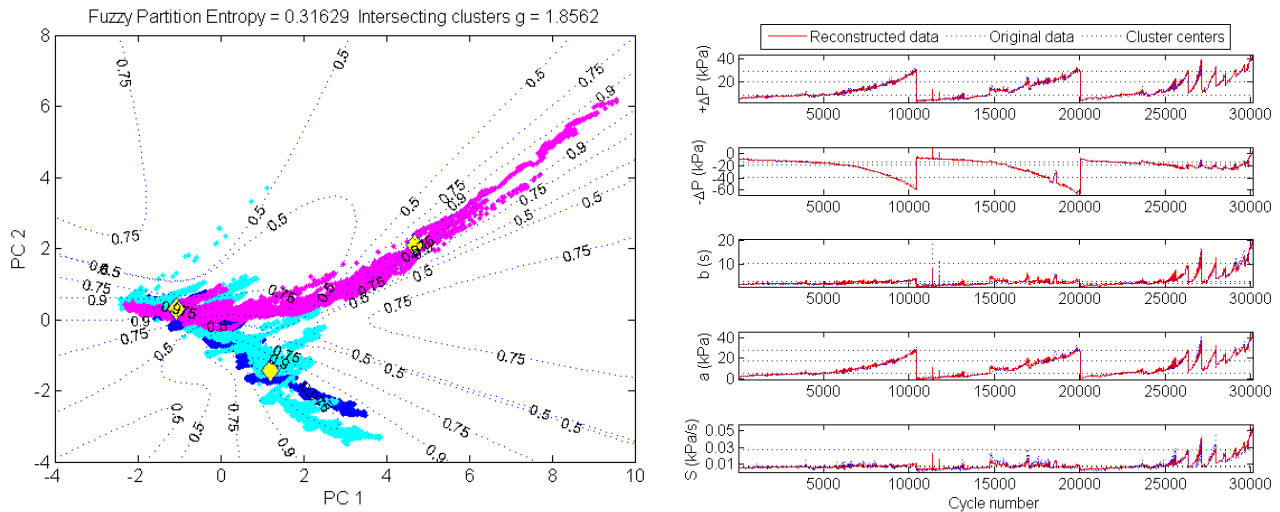


Figure 6. Left: cluster plot of the scores for the first two PCs of the long term data set. The fuzziness of the clustering can be evaluated by using the fuzzy partition entropy (0 for hard and $\ln(3)$ for fuzzy clustering), the quality by the intersecting cluster value g ($g > 1$ indicates overlapping clusters). Right: Reconstruction plot of the parameters of the long term data set. Next to the original and reconstructed data, the cluster centers are shown as horizontal dashed lines, each time in hierarchical order (clean-transitional-fouled) according to the trend of the data.

Concerning the second membrane of this data set (=M3), it was already clear from its reconstruction in the short term analysis (not shown), that S was not consistently increasing toward the fouled cluster, indicating this membrane indeed did not face high reversible fouling. Also, this membrane exhibited an aberrant behaviour if compared to the other membranes in the short term analysis (M1 and M2), while here, it acts very similar to the previously used membrane. This illustrates the dynamical character of the (dominant) fouling mechanism on longer time scales.

Pilot scale model

To check the method's applicability towards larger scaled systems, data from a pilot scale installation were analysed in a similar way. Due to the low measuring frequency (every 10 seconds compared to every second in the lab-scale case), a few problems arose. Cycle ends were not always detected successfully since the backwash length (7s) is lower than the measuring frequency. As a consequence, a lot of data were omitted by the algorithm. Next to that, the exponential part could not be approximated accurately, since at most only two data points were available. Parameters a and b were therefore excluded from further analysis. Again, two PCs were sufficient based on the 95% criterion. From the biplot (Figure 7), it can be concluded that PC_1 again gives an indication of fouling severeness, while PC_2 should still be able to discriminate reversible from irreversible fouling.

The cluster plot (Figure 7) illustrates the problem of measuring frequency by its non-qualitative clustering (both from visual inspection and the high value of g). However, certain trends can be distinguished from the patterns: PC_1 increases in time, as fouling increases. Also, a long term downward trend for PC_2 can be seen, indicating slowly increasing irreversible fouling, next to discrete upward peaks, to be seen as reversible fouling events. Nevertheless it must be stressed that clustering quality should be improved if these charts are to be used for monitoring purposes.

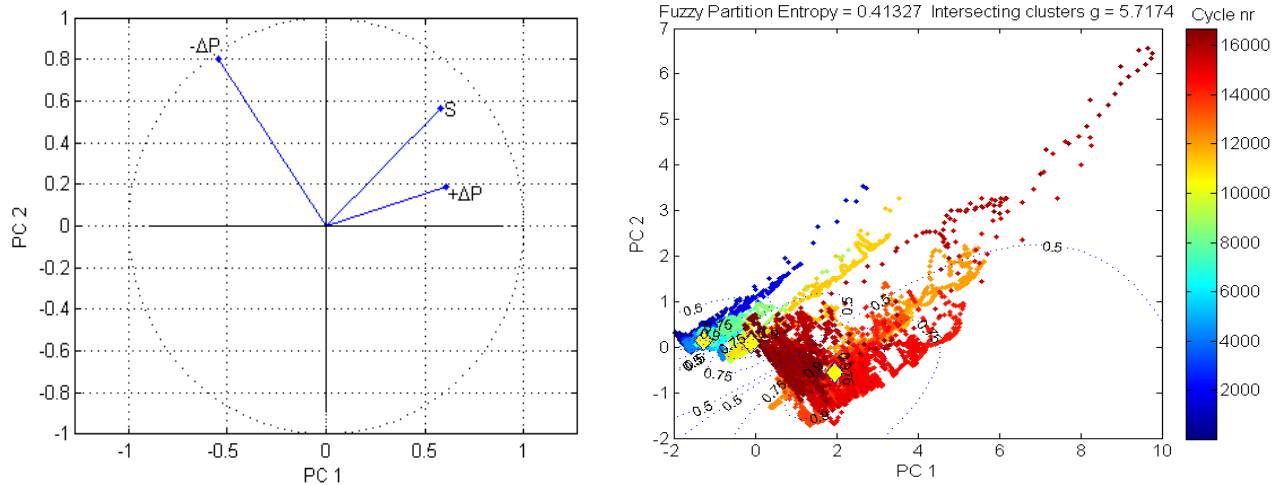


Figure 7. Left: biplot of the pilot scale installation. Right: cluster plot of the principal scores for PC_1 and PC_2 of the pilot scale installation.

FPCA versus PCA

So far, only results of the FPCA pathway have been discussed. Corresponding results of the less sophisticated PCA pathway are given here. The short term analysis yielded similar results: PC_1 is a good measure of fouling severity, while a different dominant fouling mechanism gave rise to a different model structure. Again, the hypothesis was built that for M3 irreversible fouling was dominant. Also for the long term analysis similar results were obtained, except for the location of the cluster centers, which was not as correct as for the FPCA from a visual inspection (Figure 8). For the pilot scale analysis, the low measuring frequency was responsible for a lot of scatter, the exclusion of one variable and yielded even worse clustering results (a value of 7.87 for g).

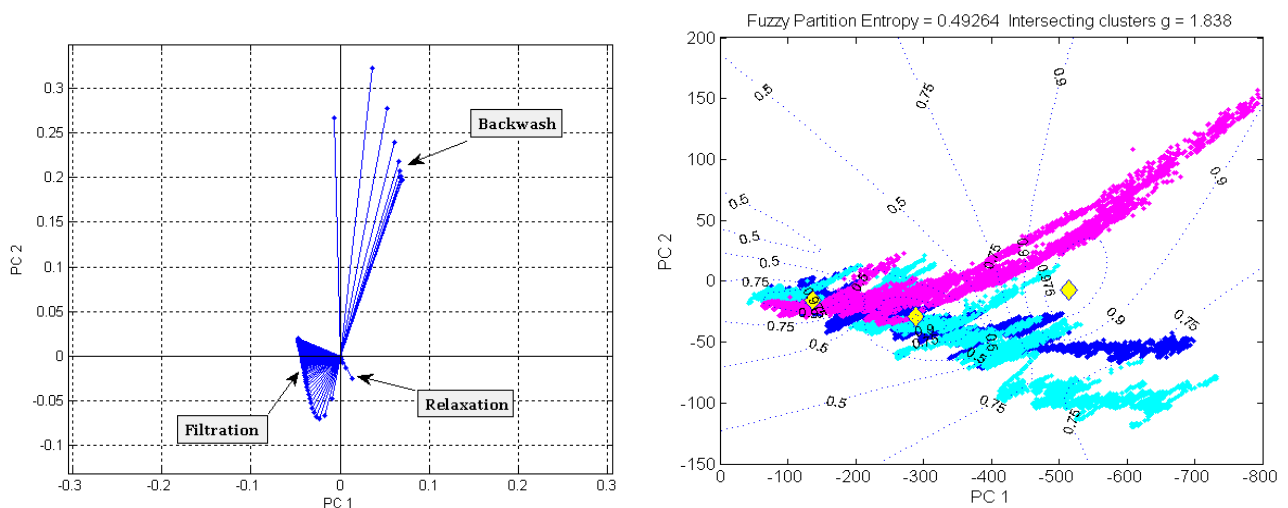


Figure 8. Left: biplot of the long term data set. Each measurement of the cycle is considered a variable, which seem to be grouped according to the cycle stage. Right: cluster plot of the principal scores of PC_1 and PC_2 of the long term PCA.

CONCLUSIONS

Principal component analysis was successfully applied to TMP data of a lab-scale installation to monitor fouling severeness, thereby relying on the first principal component. To obtain a differentiation between fouling mechanisms, the inter-membrane variance had to be included in the data set. If done so, the second principal component was able to separate reversible from irreversible fouling. With only intra-membrane variance included, the model structure adapted each time to the dominant fouling condition. The methodology was also tested on a pilot scale data set, but revealed a critical issue: the measuring frequency should be high enough to guarantee efficient data analysis. Although the first and second principal component indicated similar results to the lab-scale case, quality of the data should be increased to use this technique for real-time fouling monitoring purposes.

ACKNOWLEDGEMENT

The authors want to thank Pentair X-Flow for providing data of their pilot scale installation. Thomas Maere is supported by the Institute for Encouragement of Innovation by means of Science and Technology in Flanders (IWT).

REFERENCES

- Aguado, D. and Rosén, C. (2008). Multivariate statistical monitoring of continuous wastewater treatment plants. *Engineering Applications of Artificial Intelligence*, 21(7), 1080-1091.
- Babuska, R., Roubos, J.A. and Verbruggen H.B. (1998). Identification of MIMO systems by input-output TS fuzzy models. *Fuzzy systems proceedings, 1998. IEEE World congress on computational intelligence*, 1, 657-662.
- Baggiani, F. and Marsili-Libelli, S. (2009). Real-time fault detection and isolation in biological wastewater treatment plants. *Water Science and Technology*, 60(11), 2949-2961.
- Drews, A. (2010). Membrane fouling in membrane bioreactors – Characterisation, contradictions, causes and cures. *Journal of Membrane Science*, 363(1-2), 1-28.
- Gustafson, D.E. and Kessel, W.C. (1979). Fuzzy clustering with a fuzzy covariance matrix. *IEEE Conference on decision and control*, 17, 761-766.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(7), 498-520.
- Jiang, T., Myngheer, S., De Pauw, D.J.W., Spanjers, H., Nopens, I., Kennedy, M.D., Gary, A. and Vanrolleghem, P.A. (2008). Modelling the production and degradation of soluble microbial products (SMP) in membrane bioreactors (MBR). *Water Research*, 42(20), 4955-4964.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6), 559-572.
- Villez, K., Ruiz, M., Sin, G., Colomer, J., Rosén, C. and Vanrolleghem P.A. (2008). Combining multiway principal component analysis (MPCA) and clustering for efficient data mining of historical data sets of SBR processes. *Water Science and Technology*, 57(10), 1659-1666.